

Reduced Order Modeling and Control of Thin Film Growth in an HPCVD Reactor

S. C. Beeler¹, G. M. Kepler², H. T. Tran³, and H. T. Banks⁴

Center for Research in Scientific Computation
Department of Mathematics
North Carolina State University
Raleigh, NC 27695

Abstract. This paper describes the development of a reduced order model-based feedback control methodology for regulation of the growth of thin films in a high-pressure chemical vapor deposition (HPCVD) reactor. Precise control of the film thickness and composition is highly desirable, making real-time control of the deposition process very important. The source vapor species transport is modeled by the standard gas dynamics partial differential equations, with species decomposition reactions, reduced down to a small number of ordinary differential equations through use of the proper orthogonal decomposition technique. This system is coupled with a reduced order model of the surface reactions involved in the source vapor decomposition and film growth on the substrate. Also modeled is the real-time observation technique used to obtain a partial measurement of the deposition process.

The utilization of reduced order models greatly simplifies the mathematical formulation of the physical process so that it can be solved quickly enough to be used for real-time model-based feedback control. This control problem is fairly complicated, however, because the surface reactions render the model nonlinear. To deal with this we use a nonlinear feedback control method based on the state-dependent Riccati equation (SDRE). A second SDRE is contained in a state estimator which uses the nonlinear partial observations of the growth process to obtain an estimated state on which to base the feedback control. These nonlinear control techniques are implemented on the HPCVD model and the results analyzed as to the effectiveness of the reduced order model and nonlinear control at tracking the desired film growth profile.

¹e-mail: sbbeeler@eos.ncsu.edu

²e-mail: gmkepler@unity.ncsu.edu

³e-mail: tran@control.math.ncsu.edu

⁴e-mail: htbanks@eos.ncsu.edu

Report Documentation Page				Form Approved OMB No. 0704-0188	
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE 2000		2. REPORT TYPE		3. DATES COVERED 00-00-2000 to 00-00-2000	
4. TITLE AND SUBTITLE Reduced Order Modeling and Control of Thin Film Growth in an HPCVD Reactor				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) North Carolina State University, Center for Research in Scientific Computation, Raleigh, NC, 27695-8205				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT see report					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES 31	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

1. Introduction

Chemical vapor deposition (CVD) is a technique used to grow very thin films with certain desired properties, involving the deposition of source vapors onto a heated substrate surface where they then react chemically to form the desired material. This process is used in the manufacture of many computer hardware products, including high-speed (GaAs) integrated circuits, transistors, and DRAM chips, as well as UV detectors and green and blue light emitting diodes. A less well-known application is in high-performance electrostatic loudspeakers. Precise control of the film layer thickness and composition is extremely important, and the increasing demands on the precision of the desired properties make real-time feedback control of the CVD process very desirable [1, 2, 3, 4].

Low-pressure chemical vapor deposition processes are the preferred choice for manufacturing many of the devices mentioned above. Previous work within our N.C. State University research group has successfully implemented feedback control of film thickness and composition in GaP/Ga_{1-x}In_xP films, during experiments in a low-pressure pulsed chemical beam epitaxy (PCBE) reactor using real-time optical monitoring by *p*-polarized reflectance spectroscopy (PRS) measurements [4].

However, there are some materials (such as InN or Ga_{1-x}In_xN films) which have potential industrial uses, but cannot be effectively produced at desirable temperatures under low-pressure conditions. Extending the CVD procedure to higher pressures increases our ability to control the thermal decomposition of certain source gases, and expands the range of compositions which can be produced at optimal process temperatures. This has applications to flat panel displays covering the entire visible wavelength range, and optoelectronics in the visible to UV wavelength range, as well as radiation-resistant high power electronics. In addition, higher pressures give the advantage of a fuller ability to intentionally introduce controlled defects into the film or dope the film with impurities (for example, to give the film a positive charge, in the case of the speaker application). Control of defect chemistry/residual absorption and laser damage of nonlinear optical materials (such as ZnGeP₂) is also important for wave-guided nonlinear optical sensors and advanced optical parametric oscillators. Higher pressures can also result in faster film deposition and throughput, an advantage in time-intensive applications in the semiconductor industry. The challenge in high-pressure chemical vapor deposition (HPCVD) is that it is significantly more difficult to control than the low-pressure process, as the higher pressure introduces source vapor gas flow dynamics in the place of low-pressure ballistic source vapor pulses.

As part of a research team at N.C. State working on the design and construction of an HPCVD reactor with real-time sensors to use in feedback control of the film growth process, we have worked to create an effective mathematical model of the more complicated high-pressure deposition process. We also have developed closed-loop control methods to use on the nonlinear model, including estimation of the system state from the sensor measurements and tracking of desired properties such as film thickness and composition.

One part of the HPCVD process is the gas flow dynamics in the high-pressure reactor, as the source vapors travel from the reactor inlet to the substrate surface in a carrier gas at pressures of up to 100 atmospheres. A dilute approximation is used in our work, leading to a quasi-steady model with steady-state nonlinear continuity, momentum and energy equations being decoupled from the transient linear species equations. We use the reduced order method known as proper orthogonal decomposition (POD) to obtain a reduced order system from the species equations, so that real-time model-based feedback control is possible. In earlier work it has been shown that the POD basis

can be used to efficiently represent the species flow dynamics in an HPCVD reactor and to compute an optimal open-loop control [5], as well as to implement feedback tracking control of HPCVD gas phase species transport [6, 7]. Section 2 will describe the gas phase model, the use of the POD reduced order method, and the measurement technique used for observing the process in real time.

The second part of the CVD model is the description of the surface kinetics, including the decomposition of source vapors deposited on the surface and their reactions forming the compound which is integrated into the growing film. These reactions are represented by a reduced order surface kinetics (ROSK) model which assumes that among the many reactions involved there are a small number of significant limiting steps. The model used in this paper is a modified version of that developed in [8], coupled to the gas phase model through the flux of species to the surface. This ROSK model is described in Section 3.

The combined model is nonlinear due to the reactions on the surface, so a nonlinear feedback control method must be used. To control the growing film thickness we use a nonlinear tracking control developed in [9], which is based on the state-dependent Riccati equation (SDRE). In addition, only nonlinear partial measurements of the growth process are available. Therefore a state estimator/compensator (also from [9], using a second SDRE) is used to reconstruct from these measurements an estimated value of the full state on which to base the feedback control. In Section 4 we describe these methods in the context of the HPCVD tracking control problem.

Section 5 describes the application of the feedback control to the combined HPCVD model. Simulation results are given and analyzed, to study the effectiveness of the reduced order POD and ROSK models, the reduced order model-based nonlinear tracking control, and the state estimation process using nonlinear partial observations of the actual state. Section 6 contains some discussion of future research including theoretical considerations, and conclusions evaluating the current work are given in Section 7.

2. Reduced Order Gas-Phase Model

The gas-phase flow model for the high-pressure chemical vapor deposition reactor used here is based on the two-dimensional work in [6, 7], with a different reactor geometry and extension to three dimensions. Our model represents the flow of precursor species from the reactor inlet to the substrate surface, with the transport process dynamics given by the partial differential equations for continuity, momentum, energy and species balances, including multiple species and gas-phase reactions [10, 11, 12, 13]. These equations are considered for a case with only trace amounts of the precursors mixed with the carrier gas (N_2) in the reactor. With this dilute assumption a quasi-steady model can be constructed, with the continuity, momentum and energy equations solved as steady-state equations, based on the properties of the dominant carrier gas and independent of the reactant concentrations:

$$\vec{\nabla} \cdot (\rho \vec{v}) = 0 \quad (1)$$

$$\rho \vec{v} \cdot \vec{\nabla} \vec{v} = -\vec{\nabla} P + \vec{\nabla} \cdot \vec{\tau} - \rho \vec{g} \quad (2)$$

$$\rho c_p \vec{v} \cdot \vec{\nabla} T = \vec{\nabla} \cdot (k \vec{\nabla} T), \quad (3)$$

where the viscous stress tensor is given by

$$\vec{\tau} = -\frac{2}{3}\mu(\vec{\nabla} \cdot \vec{v})\vec{I} + \mu(\vec{\nabla} \vec{v} + \vec{\nabla} \vec{v}^T). \quad (4)$$

Here \vec{v} , T and P are the velocity, temperature and pressure, \vec{g} is the gravitational acceleration, and μ , c_p and k are the viscosity, specific heat and conductivity of the carrier gas. The density variation is modeled as

$$\rho = \rho_0[1 - \beta(T - T_0)], \quad (5)$$

with a reference temperature T_0 , a reference density ρ_0 calculated from the ideal gas law at the reference temperature and reactor pressure, and the volume coefficient of expansion $\beta = 1/T$.

We will consider a three-dimensional rectangular-box-shaped domain describing the reactor depicted in Figure 1. Previous work [6, 7] used a two-dimensional representation of an earlier HPCVD

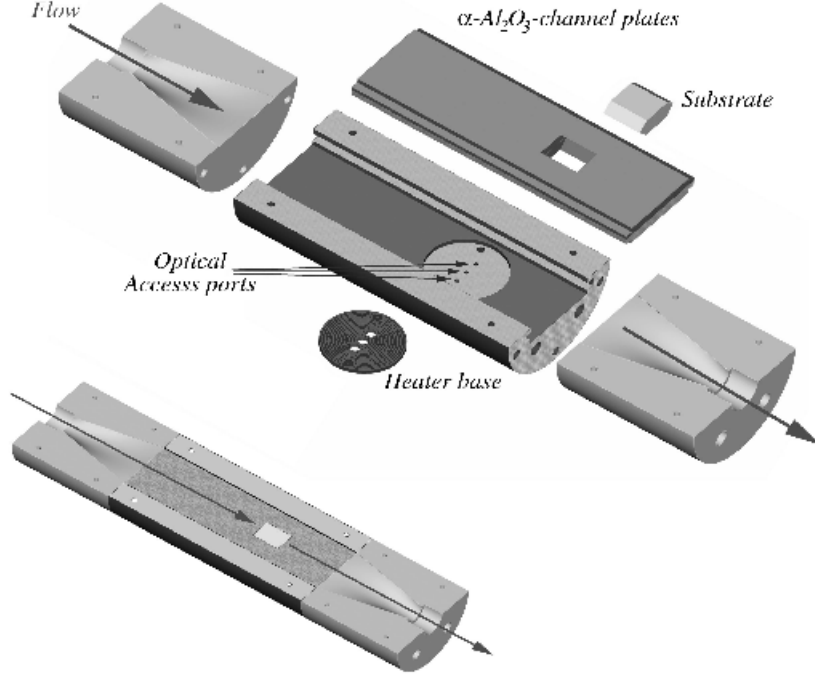


Figure 1: Three-dimensional view of the HPCVD reactor.

reactor, but all three dimensions are necessary for our new model, because in addition to the predominantly lengthwise gas flow and vertical deposition there is now also a horizontal measurement of optical absorption across the width of the reactor (which will be discussed later in this section). Our numerical model will represent an area of this reactor that is 150 mm long, 50 mm wide, and only 1 mm high. A side-view cross-section of this is shown in Figure 2, with the very small height not to scale. The portions of the reactor outside this domain, specifically the narrow inflow and outflow regions and the part of the box-shaped reactor near the inflow (seen in Figure 1) are not included in the model. There are two sapphire substrates on which the growth is to take place, each 25 mm long by 20 mm wide and placed in the center of the reactor model (width-wise and length-wise), located symmetrically on both the top and bottom of the flow channel. The growth is driven thermally by a circular heating element, or susceptor, 40 mm in diameter and centered behind each substrate. The substrate and susceptor placement can be seen in Figures 1 and 2. The gas flow enters through the inlet seen at the left end of the reactor in these figures, flows across the

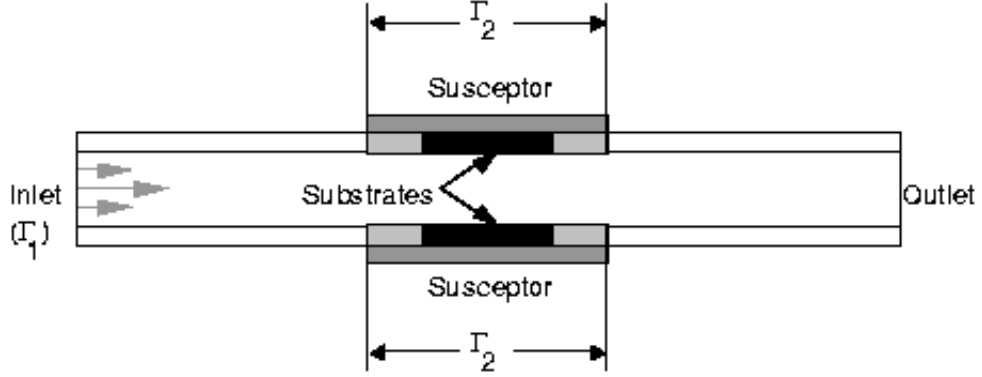


Figure 2: HPCVD reactor side-view cross-section (not to scale).

substrate surfaces (depositing some of the source species there), and exits the outlet at the right end of the reactor. Here we consider Nitrogen (N_2) as the carrier gas, at a pressure of 10 atmospheres. Temperature-dependent values of μ , c_p and k for N_2 are linearly interpolated from values in the literature [14, 15, 16].

The boundary conditions for the steady-state flow model will be formulated as follows. There is a no-slip (zero velocity) condition on all walls of the reactor as well as the substrate, while the inflow velocity profile corresponds to rectangular duct flow [17] with a flow rate of 10 standard liters per minute (slpm). The temperature boundary condition is room temperature (298 K) at the inlet and 1000 K at the susceptors. Boundary conditions on the temperature of the reactor walls include conduction along the walls, radiation off the walls, convective heat loss to the environment, and modeled heat loss to the outside of the reactor. A special boundary condition in the commercial software package FIDAP (Fluid Dynamics International, Evanston, IL), described in [12], page 6-14, is used to obtain a smooth outflow condition. The pressure is involved in equations (1)-(5) only as $\vec{\nabla}P$, so therefore only pressure differences matter, not absolute pressures, and an implicit $P = 0$ boundary condition is used at the outlet. A Galerkin finite element method with weighted residuals for the degrees of freedom (\vec{v} , T and P) is used to discretize the gas flow equations. The discretization uses a mixed formulation with 4983 brick elements (corresponding to 47131 nodes), with piecewise linear discontinuous elements for pressure and quadratic (27-noded) elements for the other degrees of freedom. We used FIDAP to solve this finite element problem for steady-state values of \vec{v} , T and P at the nodal points.

The solutions for velocity and temperature that are found from equations (1)-(5) can then be used in the solution of the time-dependent species equations for the precursor mass fractions,

$$\frac{\partial Y_n}{\partial t} + \vec{v} \cdot \vec{\nabla} Y_n = \frac{1}{\rho} \vec{\nabla} \cdot (\rho D_n \vec{\nabla} Y_n) + \sum_{i=1}^{N_R} r_{ni}, \quad (6)$$

where Y_n is the mass fraction of the n th species, D_n is the diffusivity of species n , N_R is the number of gas phase reactions, and r_{ni} is the rate of production of species n in the i th reaction.

In the HPCVD computational experiments to follow we will consider trimethylgallium, or TMG ($Ga(CH_3)_3$), and phosphine (PH_3) as source materials for the growth of the III-V film gallium phosphide (GaP). For now we only investigate using one TMG pulse to control its contribution

to the film deposition process. This can be done since the inputs are pulsed with pauses between them, so that the gallium species and phosphorus species are not present in the reactor at the same time (though a constant overall gas flow profile is maintained throughout the pulsing cycle by the carrier gas). This pulsing prevents nucleation of GaP material in the gas phase. For the gallium transport we will consider three gas-phase species (aside from the carrier gas): Y_1 representing the mass fraction of TMG, Y_2 that of dimethylgallium (DMG), and Y_3 that of monomethylgallium (MMG).

We will assume that there are $N_R = 2$ significant gas phase reaction mechanisms for the gallium species. There is the reaction for the decomposition of TMG to DMG and a methyl molecule, $\text{Ga}(\text{CH}_3)_3 \rightarrow \text{Ga}(\text{CH}_3)_2 + \text{CH}_3$ (reaction 1), and the decomposition of DMG to MMG and a methyl molecule, $\text{Ga}(\text{CH}_3)_2 \rightarrow \text{GaCH}_3 + \text{CH}_3$ (reaction 2). These decompositions can be described as first-order Arrhenius reactions with rates of production given by

$$r_{ni} = \nu_{ni} \frac{W_n}{W_{m_i}} k_i e^{-E_i/RT} Y_{m_i}, \quad (7)$$

where m_i is the number of the species which is the source in reaction i . The parameter ν_{ni} is the stoichiometric constant for species n in reaction i , W_n and W_{m_i} are the molecular weights of those particular species, and k_i is the rate constant and E_i the activation energy for reaction i . The values of ν_{ni} , k_i and E_i for the two reactions we consider are taken from [18, 19] and are given in Table 1.

Reaction i	Species n	$\nu_{ni} k_i$ (s^{-1})	E_i (kcal/mol)
1	1 (m_1)	-5.5×10^{15}	61.0
1	2	5.5×10^{15}	61.0
2	2 (m_2)	-8.7×10^7	35.4
2	3	8.7×10^7	35.4

Table 1: Rate constants and activation energies for gas-phase reactions.

The molar weights for the three species are $W_{TMG} = 114.8$ g/mol, $W_{DMG} = 99.79$ g/mol, and $W_{MMG} = 84.755$ g/mol, and $R = 1.99$ cal/(mol·K) is the universal gas constant. The temperature-dependent values of the diffusivities D_n are linearly interpolated from values in the literature [20]. The methyl molecules do not participate in the film growth and, due to the dilute approximation, do not contribute to the flow properties, so we do not include them in the transport equations.

The boundary conditions for the species transport equations will be as follows. We consider all of the area in contact with the two susceptors (including both the substrates and some of the surrounding reactor wall area) to be perfectly absorbing for each species (this area is shown as Γ_2 in Figure 2). The remainder of the reactor wall areas are considered to be completely non-absorbing, also for each species. At the inlet (Γ_1 in Figure 2), the source species (TMG) is set to desired values, and will be used as the input for control of the growth process, while the other two species are set to 0. Zero normal flux is assumed at the outflow, so that $\partial Y_n / \partial \vec{n} = 0$. The initial condition of the system is that no species are present in the reactor. Thus the entire time-dependent species transport system based on equation (6) is given by

$$\begin{aligned}
\frac{\partial Y_n}{\partial t} + \vec{v} \cdot \vec{\nabla} Y_n &= \frac{1}{\rho} \vec{\nabla} \cdot (\rho D_n \vec{\nabla} Y_n) + \sum_{i=1}^{N_R} r_{ni} \\
Y_n(0, \vec{x}) &= 0 \\
Y_1(t, \vec{x}) &= u_1 \quad \text{at inflow } \Gamma_1 \\
Y_n(t, \vec{x}) &= 0 \quad \text{at inflow } \Gamma_1 \quad (n = 2, 3) \\
Y_n(t, \vec{x}) &= 0 \quad \text{at substrates, susceptors } \Gamma_2 \\
\frac{\partial Y_n(t, \vec{x})}{\partial \vec{n}} &= 0 \quad \text{at walls, outflow,}
\end{aligned} \tag{8}$$

where r_{ni} are the reaction production rates from equation (7) and u_1 is the control input which will be discussed later.

We use a penalty boundary formulation on the species transport problem in order to change all the boundary conditions to Neumann conditions. This prepares the system for the Galerkin procedure to be discussed later in this section and the control problem to be described in Section 4. Under the penalty boundary formulation, the system (8) is modified to become

$$\begin{aligned}
\frac{\partial Y_n}{\partial t} + \vec{v} \cdot \vec{\nabla} Y_n &= \frac{1}{\rho} \vec{\nabla} \cdot (\rho D_n \vec{\nabla} Y_n) + \sum_{i=1}^{N_R} r_{ni} \\
Y_n(0, \vec{x}) &= 0 \\
\frac{\partial Y_1(t, \vec{x})}{\partial \vec{n}} &= \frac{1}{\epsilon} (Y_1(t, \vec{x}) - u_1) \quad \text{at inflow } \Gamma_1 \\
\frac{\partial Y_n(t, \vec{x})}{\partial \vec{n}} &= \frac{1}{\epsilon} Y_n(t, \vec{x}) \quad \text{at inflow } \Gamma_1 \quad (n = 2, 3) \\
\frac{\partial Y_n(t, \vec{x})}{\partial \vec{n}} &= \frac{1}{\epsilon} Y_n(t, \vec{x}) \quad \text{at substrates, susceptors } \Gamma_2 \\
\frac{\partial Y_n(t, \vec{x})}{\partial \vec{n}} &= 0 \quad \text{at walls, outflow,}
\end{aligned} \tag{9}$$

where ϵ is a small parameter (for our simulations we use $\epsilon = 10^{-3}$). With sufficient regularity assumed, it can be argued that as $\epsilon \rightarrow 0$ the solution to (9) approximates the solution to the problem in (8) with $Y_1(t, \vec{x}) = u_1$ and $Y_{2,3}(t, \vec{x}) = 0$ at the inlet and $Y_n(t, \vec{x}) = 0$ at the substrate (see [21, 22] for related discussions). This penalty formulation was successfully implemented in our previous feedback controlled HPCVD simulations for different reactor geometries in [6, 7].

The Proper Orthogonal Decomposition (POD), also known as the Karhunen-Loève expansion [23, 24] or principal component analysis [25], is a method of choosing a set of basis functions tailored to a particular problem so that a minimal number of functions are needed to represent that problem's dynamics. This is a well-known method of feature extraction in the pattern recognition field [26], and it has been used on many types of physical problems, including turbulent flows [27, 28, 29, 30, 31]. It has been used in open-loop control of CVD systems [5, 32], and recently in closed-loop control of CVD systems in [6, 7], for discretizing species transport PDEs similar to system (9) so that they may be numerically solved more quickly. In contrast to the POD method, a standard finite difference, finite element, or spectral method of discretization will result in a very large system of ordinary differential equations. The POD method utilizes observations of the problem dynamics to choose

basis functions which include as much of those dynamics as possible, thus potentially making it possible to reduce the number of approximate ODEs needed from hundreds or thousands down to the order of ten or less. The most important consequence of this in our work is that the POD system is small enough to use to generate a real-time on-line feedback control, whereas the system resulting from a generic discretization method would require too much time to solve.

In our problem we use FIDAP simulations of the species flow equations described above to obtain a set of experimental observations, or “snapshots”, of the species mass fractions at various times, to serve as the raw material for finding the POD modes. The snapshot data set for the n th species consists of K vectors, $U_n = \{u_{n1}^N, u_{n2}^N, \dots, u_{nK}^N\}$, each vector consisting of the N nodal values of the mass fraction of the n th species at a different time t_i ($i = 1, 2, \dots, K$) during the simulation. This simulation is done starting with zero mass fraction for each species in the body of the reactor, with a constant TMG mass fraction input serving as a nontrivial “control” for 0.5 s, followed by a clearing pulse of zero input for another 0.5 s. The simulation uses a backward Euler method with time steps of 0.005 s or 0.01 s for the time integration. A total of 100 snapshots are obtained at 0.01 s time intervals.

For the purposes of describing the process of finding the POD, we will think of the snapshots $\{u_{ni}^N\}$ not as length N vectors but as functions $\{u_{ni}^N(x)\}$ of the location x in the reactor domain Ω . We also will eliminate the species subscript n and length- N superscript for now, so that the set of snapshots is given by $U(x) = \{u_1(x), u_2(x), \dots, u_K(x)\}$, remembering that we are considering the snapshots for a single species. Following the description of [5], we look for the set of K POD modes that most closely match the snapshot data, so that the POD modes $z = z(x)$ maximize

$$\frac{1}{K} \sum_{i=1}^K |(u_i, z)|^2, \quad (10)$$

while still being independent of each other and having

$$(z, z) = \|z\|^2 = 1.$$

Here (\cdot, \cdot) and $\|\cdot\|$ are the \mathbf{L}^2 inner product and norm. Specifically we look for POD modes that are linear combinations of the snapshots, so that

$$z(x) = \sum_{i=1}^K a_i u_i(x), \quad (11)$$

with the coefficients a_i to be found so that (10) is maximized.

To start, we define the function

$$\mathbf{K}(x, x') = \frac{1}{K} \sum_{i=1}^K u_i(x) u_i(x'), \quad (12)$$

and the operator \mathbf{R} on z

$$(\mathbf{R}z)(x) = \int_{\Omega} \mathbf{K}(x, x') z(x') dx'. \quad (13)$$

By manipulating these defined equations we find that

$$(\mathbf{R}z, z) = \int_{\Omega} \mathbf{R}z(x) z(x) dx$$

$$\begin{aligned}
&= \int_{\Omega} \int_{\Omega} \mathbf{K}(x, x') z(x') dx' z(x) dx \\
&= \frac{1}{K} \sum_{i=1}^K \int_{\Omega} \int_{\Omega} u_i(x) u_i(x') z(x') dx' z(x) dx \\
&= \frac{1}{K} \sum_{i=1}^K |(u_i, z)|^2.
\end{aligned} \tag{14}$$

It also follows that $(\mathbf{R}z_1, z_2) = (z_1, \mathbf{R}z_2)$ for any $z_1, z_2 \in \mathbf{L}^2(\Omega)$, so therefore \mathbf{R} is a nonnegative symmetric operator on $\mathbf{L}^2(\Omega)$. With this property, the problem of maximizing the expression (10) is the same as finding the largest eigenvalue in the problem

$$\mathbf{R}z = \lambda z \tag{15}$$

subject to $\|z\|^2 = 1$, or equivalently,

$$\int_{\Omega} \mathbf{K}(x, x') z(x') dx' = \lambda z(x) \tag{16}$$

with $\|z\|^2 = 1$.

By substituting the definitions of z and \mathbf{K} in equations (11) and (12) into equation (16), we obtain

$$\sum_{i=1}^K \left[\sum_{j=1}^K \frac{1}{K} \int_{\Omega} u_i(x') u_j(x') dx' a_j \right] u_i(x) = \sum_{i=1}^K \lambda a_i u_i(x), \tag{17}$$

which can be rewritten as the matrix eigenvalue problem

$$C\psi = \lambda\psi, \tag{18}$$

where the matrix and eigenvector are given by

$$C_{ij} = \frac{1}{K} \int_{\Omega} u_i(x') u_j(x') dx', \tag{19}$$

$$\psi = [a_1, a_2, \dots, a_K]^T. \tag{20}$$

C is a nonnegative Hermitian matrix, so therefore it has a complete set of orthogonal eigenvectors $\{\psi_1, \psi_2, \dots, \psi_K\}$, with $\psi_k = [a_1^k, a_2^k, \dots, a_K^k]^T$, ordered from largest to smallest eigenvalues. The expression (10) is then maximized by the POD mode obtained from the eigenvector corresponding to the largest eigenvalue by setting

$$z_1(x) = \sum_{i=1}^K a_i^1 u_i(x) = U(x) \psi_1. \tag{21}$$

The other POD modes are calculated similarly from the remaining $K - 1$ eigenvectors.

In the case of our HPCVD problem, the snapshots are actually given in terms of length N vectors of the finite element coefficients, so instead of the integral version of the correlation matrix C , it will be given as $C_{ij} = (1/K) u_i^T u_j$. For a particular species n , the i th POD mode z_{ni}^N in the set $Z_n = \{z_{n1}^N, z_{n2}^N, \dots, z_{nK}^N\}$ is obtained by taking a linear combination of the snapshots $U_n = \{u_{n1}^N, u_{n2}^N, \dots, u_{nK}^N\}$

of that species. That is, $z_{ni}^N = U_n \psi_{ni}^K$, where $\{\psi_{n1}^K, \psi_{n2}^K, \dots, \psi_{nK}^K\}$ are the orthogonal eigenvectors of the correlation matrix of the snapshots. These are given by the solution of $\left[(1/K)U_n^T U_n\right] \psi_{ni}^K = \lambda_{ni} \psi_{ni}^K$, ranked in descending order of the eigenvalues λ_{ni} . The POD basis elements are found from these POD modes by using generic finite element interpolation functions, such as quadratic functions. A Galerkin procedure on the weak form of the original PDE system using a chosen number of these basis elements will result in an ODE system similar to that of a finite element Galerkin procedure but with significantly fewer equations.

The POD basis has certain desirable properties which result from the conditions of its formulation. First, the POD modes can be shown to be orthonormal (see e.g. [5]). Secondly, the POD coefficients are uncorrelated (see either [28] page 77 or [29] page 237). Finally, POD has the property that the representation of the original data set $U = \{u_i(x)\}$ in terms of the first M POD modes,

$$\{u_i(x)\} \cong \left\{ \sum_{j=1}^M b_{ij} z_j(x) \right\}, \quad (22)$$

for any $M \leq K$, maximizes the data variability of such an M -mode reduced basis representation of the snapshots. That is, for a representation in terms of any other orthonormal basis,

$$\{u_i(x)\} \cong \left\{ \sum_{j=1}^M c_{ij} \tilde{z}_j(x) \right\}, \quad (23)$$

the POD coefficients will contain more of the variation in the data set, as measured by $\sum_{i=1}^M \langle b_i^2 \rangle$ using the notation $\langle b_i b_j \rangle = (1/K) \sum_{k=1}^K b_{ki} b_{kj}$, than the coefficients $\{c_{ij}\}$ for the other basis. More specifically, it can be shown that

$$\sum_{i=1}^M \langle b_i^2 \rangle = \sum_{i=1}^M \lambda_i \geq \sum_{i=1}^M \langle c_i^2 \rangle \quad (24)$$

(also see [28] page 77 or [29] page 237). The “energy” of the POD modes described by Berkooz in these places is what we refer to here as the data variability. In an analogous concept, it can be shown that the POD basis minimizes the mean square error of an M -mode representation over all choices of orthonormal bases [26].

As a consequence of the property described above, the percentage of the total snapshot set data variability contained in a certain POD mode z_k is given by the ratio of the eigenvalue λ_k to the total of all eigenvalues, $\lambda_k / \sum_{j=1}^K \lambda_j$. The reason for ordering the POD modes from highest to lowest eigenvalues is to include as much of the variability of the system into the few first modes as possible. Therefore a POD reduced order system, when found with a Galerkin procedure using these ordered modes, can still be a very good representation of the dynamics of the system while using only the first few modes. Due to these characteristics, one approach to choosing the number of POD modes is to choose the smallest M so that the following condition is satisfied:

$$\sum_{i=1}^M \lambda_i / \sum_{j=1}^K \lambda_j > 0.99. \quad (25)$$

The reduced order gas-phase model is produced from a discretization of the flow problem (9) using the POD modes in the following way. This model uses only a small number of basis functions

and will be used to find the real-time feedback control. We start with the weak formulation of the species flow equation:

$$\begin{aligned} \int_{\Omega} \frac{\partial Y_n}{\partial t} w_j d\Omega &= - \int_{\Omega} (\vec{v} \cdot \vec{\nabla} Y_n) w_j d\Omega - \int_{\Omega} D_n \vec{\nabla} Y_n \cdot \vec{\nabla} w_j d\Omega + \int_{\Omega} \frac{1}{\rho} D_n (\vec{\nabla} Y_n \cdot \vec{\nabla} \rho) w_j d\Omega \\ &+ \sum_{i=1}^{N_R} \int_{\Omega} r_{ni} w_j d\Omega + \frac{1}{\epsilon} \int_{\Gamma_1, \Gamma_2} w_j D_n Y_n ds - \frac{1}{\epsilon} \int_{\Gamma_1} w_j D_n u_n ds, \end{aligned} \quad (26)$$

for $n = 1, 2, 3$, with $u_n = 0$ for $n = 2, 3$ (the only species inlet flow we control is TMG). The important portions of the boundary involved in this formulation are Γ_1 , which represents the inlet, and Γ_2 , which represents the substrates and susceptors (these boundary areas can be seen in Figure 2). The values of \vec{v} and T , and by extension $\rho(T)$ and $D_n(T)$, are found from the steady-state flow simulations, and the reaction rates of production r_{ni} are found by equation (7) in terms of T and \vec{Y} .

We then form the POD basis elements from the calculated POD modes $\{z_{nk}^N\}$ and the finite element quadratic interpolation functions $\{\phi_k\}$. The formula

$$\Phi_{nk}(\vec{x}) = \sum_{i=1}^N (z_{nk}^N)_i \phi_i(\vec{x}) \quad (27)$$

gives a set of K basis functions $\{\Phi_{nk}\}$ for each species $n = 1, 2, 3$. The mass fraction of the n th species is approximated as a linear combination of the M_n most significant of these POD basis elements for that species,

$$Y_n^{M_n}(t, \vec{x}) = \sum_{k=1}^{M_n} y_{nk}(t) \Phi_{nk}(\vec{x}), \quad (28)$$

where $M_n \ll K \ll N$. Substitution of this approximation into (26), using test functions $w_j = \Phi_{nj}$ ($j = 1, \dots, M_n$) for the n th species, results in the reduced order system

$$\dot{y}^{M_G}(t) = A^{M_G} y^{M_G}(t) + B^{M_G} u(t), \quad (29)$$

where $M_G = \sum_{n=1}^3 M_n$, A^{M_G} is an $M_G \times M_G$ matrix and B^{M_G} is an $M_G \times 1$ vector. This POD discretization produces an ODE system in terms of the POD coefficients $\{y_{nk}\}$, in a form appropriate for a feedback control implementation and of sufficiently small order that real-time model-based control is possible. The POD snapshots are obtained, as discussed above, using FIDAP simulations of the flow problem (9). These use a similar discretization process to that just described, but with a very large number of finite element basis functions rather than the small number of specialized POD basis functions.

The reactor model we use here is required to be three-dimensional because the real-time observation of the gas-phase species transport system is done with a measurement of the intensity of a beam of light traveling across the width of the reactor (the dimension previously eliminated in the two-dimensional reactor calculations of [6, 7]) above a substrate surface. By using a three-dimensional simulation we also create a more realistic representation of the actual flow dynamics in the physical reactor. The amount of light reaching the detector after absorption in the reactor depends on the light beam frequency and the gas-phase species mass fractions along the path of the beam. The light source and light detection locations on the sides of the reactor for this horizontal observation technique are shown in Figure 3. The optical access ports behind the substrates are for a reflectance

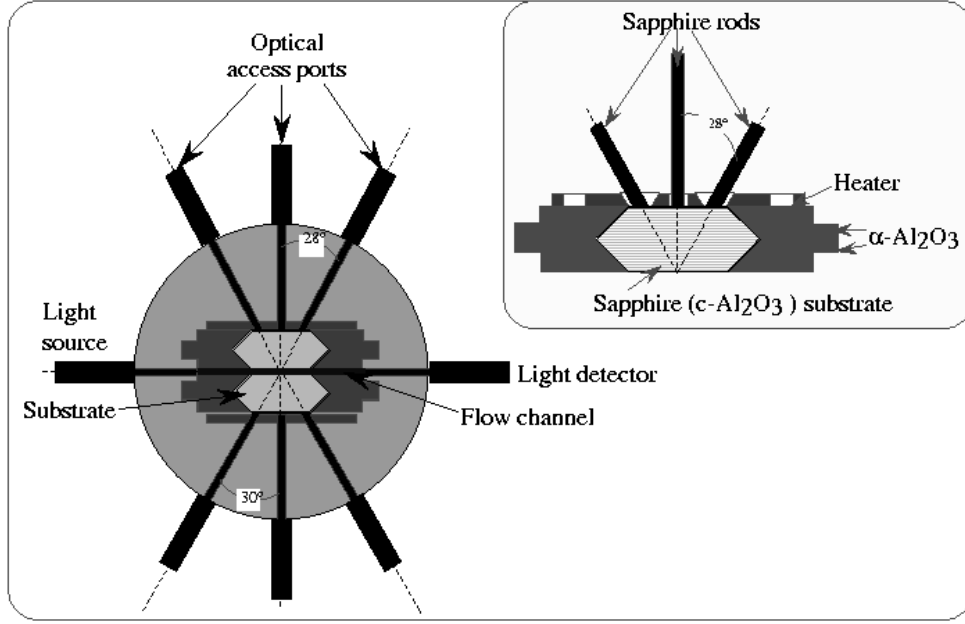


Figure 3: Measurement techniques in the HPCVD reactor. In this figure the primary gas flow direction is perpendicular to the page.

measurement taken from the reverse side of the substrates, using the amount of light reflected back to obtain a partial measurement of the properties of the film and surface reaction layer on the substrate surface. Our current model does not include this as an observation method (see [4] for control problems using related PRS sensors). We include only the optical absorption measurement of the gas-phase species in the present initial proof-of-concept calculations for this model.

The intensity of the light signal as measured at the detector side of the channel is

$$I = I_0 \exp \left(- \int_{\mathcal{W}} \frac{4\pi}{\lambda} \epsilon_{imag}(\vec{x}) d\vec{x} \right), \quad (30)$$

where \mathcal{W} represents the path across the 50 mm width of the reactor above the target point \vec{x}_p at the center of one of the substrates. The wavelength of the light is λ , and I_0 and I are the initial and detected intensities respectively. The parameter $\epsilon_{imag}(\vec{x})$ is the imaginary component of the dielectric function of the gas mixture at a location \vec{x} in the reactor. This dielectric function is given by

$$\begin{aligned} \epsilon(\vec{x}) &= \epsilon_{real}(\vec{x}) - i\epsilon_{imag}(\vec{x}) \\ &= 1 + \left(\sum_{k=0}^N \frac{Y_k(\vec{x})}{W_k} \right)^{-1} \sum_{n=0}^N \frac{Y_n(\vec{x}) F_n}{W_n}, \end{aligned} \quad (31)$$

where N is the number of gas-phase species ($N = 3$ in our model) and species $n = 0$ represents the carrier gas. Y_n is the mass fraction of species n in the gas-phase model, and W_n is the molar weight of species n . The parameter $F_n = a_n - ib_n$ is the complex optical response of species n .

The wavelength λ is typically chosen so that the absorption measurement is particularly sensitive to one species. This is achieved through using a wavelength at which the imaginary part of the optical

response of that particular species is larger than those of the other species. Since experimental data is not yet available, we use the following assumed values for the sake of implementation: $b_0 = 0$, $b_1 = 5 \times 10^{-3}$, and $b_2 = b_3 = 5 \times 10^{-4}$, all at $\lambda = 1000$ nm. With these values and the assumption of a dominant carrier gas due to the dilute approximation, the dielectric function imaginary part is given by

$$\epsilon_{imag}(\vec{x}) = W_0 \sum_{n=1}^3 \frac{b_n Y_n(\vec{x})}{W_n}. \quad (32)$$

The molar weight values are $W_0 = 28.01$ g/mol, $W_1 = 114.8$ g/mol, $W_2 = 99.79$ g/mol, and $W_3 = 84.755$ g/mol. To normalize the observed signal described by equations (30) and (32), we consider the actual observation, which will be used in the state estimation process described in Section 4, to be

$$\begin{aligned} z(\vec{Y}) &= 1 - \frac{I}{I_0} \\ &= 1 - \exp \left(-\frac{4\pi W_0}{\lambda} \int_{\mathcal{W}} \sum_{n=1}^3 \frac{b_n}{W_n} Y_n(\vec{x}) d\vec{x} \right). \end{aligned} \quad (33)$$

3. Surface Kinetics Model

The gas-phase reduced order model described in the last section will be linked with a modified version of the reduced order surface kinetics (ROSK) model developed in [8]. The model here is slightly altered since we are only considering one TMG source pulse, so the surface model described below will only have a single source term based on the flux at the surface from the gallium-containing species in the gas phase. From the species mass fractions Y_n , the flux of gallium to a specific point \vec{x}_p at the center of one of the substrate surfaces is given by

$$q(\vec{Y}) = -\rho \left[D_1 \frac{W_{Ga}}{W_1} \frac{\partial Y_1}{\partial \vec{n}} \Big|_{\vec{x}_p} + D_2 \frac{W_{Ga}}{W_2} \frac{\partial Y_2}{\partial \vec{n}} \Big|_{\vec{x}_p} + D_3 \frac{W_{Ga}}{W_3} \frac{\partial Y_3}{\partial \vec{n}} \Big|_{\vec{x}_p} \right]. \quad (34)$$

Writing this gallium flux as an approximation in terms of the gas-phase POD system (29), using the representation of the state in equation (28), we have

$$\begin{aligned} q^{MG}(y^{MG}(t)) &= -\rho \sum_{n=1}^3 \left[D_n \frac{W_{Ga}}{W_n} \sum_{k=1}^{M_n} \frac{\partial \Phi_{nk}}{\partial \vec{n}} \Big|_{\vec{x}_p} y_{nk}(t) \right] \\ &= H_q^{MG} y^{MG}(t). \end{aligned} \quad (35)$$

The molar weight values are $W_1 = 114.8$ g/mol, $W_2 = 99.79$ g/mol, $W_3 = 84.755$ g/mol, and $W_{Ga} = 69.9$ g/mol. The density ρ and diffusivities D_n are dependent on the temperature at the point \vec{x}_p on the substrate, with the density value found through equation (5) and the diffusivity values interpolated from literature values [20]. This flux will be used to couple the gas phase state variables (the POD coefficients y_{nk}) to the surface model, which has state variables n_i representing the concentrations of species in the film and the surface reaction layer (SRL) on top of the growing film.

Since there is substantial decomposition during the flow through the high-pressure reactor, we will consider n_1 , the intermediate stage of the gallium species decomposition, to represent the amount

of gallium in the SRL. The species n_2 will represent the amount of activated gallium (available to react, unlike n_1) in the SRL, while n_3 will represent the amount of gallium phosphide in the film. Since we are looking only at a single TMG source pulse we will consider the phosphorus source to have already arrived at the surface. The initial concentration of activated phosphorus in the SRL is chosen as an amount in excess of that needed to grow the desired GaP layer thickness, and is specifically given by $S_p = \gamma S_0 / N_A$, where S_0 is the density of surface sites on the substrate, N_A is Avogadro's number, and $\gamma > 1$ is a chosen constant. The concentration of activated surface phosphorus at any time can be found from the initial amount on the surface less the amount incorporated into the film up to that time, as $S_p - n_3(t)$, thus removing the need for an ODE representing activated phosphorus. This serves to simplify the model, also making it more easily controllable, since the amount of activated surface phosphorus would be very difficult to regulate without any phosphine input control. The model is also simplified by removing the desorption loss terms present in [8], as they should be less significant under the high surface pressure in the HPCVD reactor than in the near-vacuum reactor used in [8].

In this model the variables n_i give the molar concentrations of the surface species at the particular point \vec{x}_p at the center of a substrate. The gas-phase flux is used as the source term for gallium in the surface reaction layer by the formula $S_1(t) = q^{M_G}(t) / W_{Ga}$. All the state variables in the surface model ODE system are assumed to be initially zero. With these properties, the modified ROSK model is given by

$$\dot{n}_1(t) = \frac{q^{M_G}(t)}{W_{Ga}} - k_1 n_1(t) \quad (36)$$

$$\dot{n}_2(t) = k_1 n_1(t) - k_{GaP} [S_p - n_3(t)] n_2(t) \quad (37)$$

$$\dot{n}_3(t) = k_{GaP} [S_p - n_3(t)] n_2(t), \quad (38)$$

with n_1 , n_2 , and n_3 the moles per m^2 of Ga, activated Ga, and GaP respectively. For this modified ROSK model we have chosen the rate constants to be $k_1 = 20 \text{ s}^{-1}$ (for transformation from gallium to activated gallium) and $k_{GaP} = 2000 \text{ m}^2/(\text{mol}\cdot\text{s})$ (for formation of gallium phosphide). Values of these parameters which better fit the physical surface processes should be found from experimental data as in [8], but our chosen values will serve for a proof-of-concept demonstration of the model and its control. Also, more detail may need to be added to the model in the future, in the form of new terms and/or state variables, once experimental data from the reactor is obtained and studied. The surface model has size $M_S = 3$ and can be written in vector form as $\dot{n}^{M_S} = f(n^{M_S}, q^{M_G})$. The combined gas-phase/surface system will have size $M = M_G + M_S$.

4. Feedback Control Formulation

Having constructed a reduced order model of the HPCVD process, we wish to implement a feedback control in order to force the film growth to follow a certain profile. The combined system, consisting of the gas phase model found in Section 2 linked to the modified ROSK model described in Section 3, is given by

$$\begin{bmatrix} \dot{y}^{M_G}(t) \\ \dot{n}^{M_S}(t) \end{bmatrix} = \begin{bmatrix} A^{M_G} y^{M_G}(t) \\ f(n^{M_S}(t), q^{M_G}(t)) \end{bmatrix} + \begin{bmatrix} B^{M_G} \\ 0 \end{bmatrix} u(t). \quad (39)$$

This can be written in terms of a single state vector as

$$\dot{x}^M(t) = \mathcal{A}(x^M(t))x^M(t) + \mathcal{B}u(t), \quad (40)$$

with $\mathcal{A}(x^M)$ an $M \times M$ matrix function of the combined state x^M , \mathcal{B} a length M constant-valued vector, and u the single control input (the inlet TMG mass fraction). The initial state x_0^M is zero for all the gas-phase coefficients as well as all the surface species. We will track the gallium phosphide film thickness

$$\begin{aligned} d^M(t) &= V_{GaP} n_3(t) \\ &= H^M x^M(t), \end{aligned} \quad (41)$$

where $V_{GaP} = 12.2 \text{ cm}^3/\text{mol}$ is the molar volume of GaP.

To construct the tracking control problem for the combined system (40) and tracking signal (41) we start with the cost functional

$$J(x_0^M, u) = \frac{1}{2} \int_0^\infty \left[(d^M - d_T)^T Q_1 (d^M - d_T) + (\bar{d}^M)^T Q_2 \bar{d}^M + u^T R u \right] dt. \quad (42)$$

The weights are $R = 1$, $Q_1 = r_1$, and $Q_2 = r_2 I_M$, with I_M the $M \times M$ identity matrix and r_1 and r_2 two chosen design parameters. The desired film thickness trajectory $d_T(t)$ is a smoothed upward ramp function of growth from 0 to the final thickness of one monolayer of GaP, roughly 10^{-11} m (see Figure 4). The time axis in this figure has been nondimensionalized as well; each

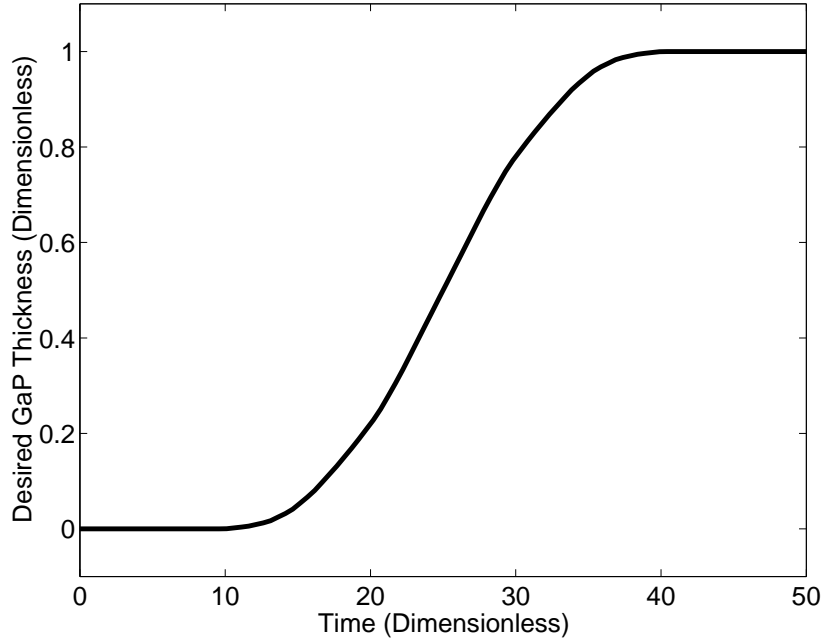


Figure 4: Desired film thickness growth profile.

time unit corresponds to 0.1 s. In addition to the weight on the tracking signal d^M and the weight on the control u there is also a weight on the other state variables given by \bar{d}^M , in order to minimize any undesirable behavior in the system not necessary to tracking the chosen thickness profile. The method of and reasoning behind adding this term are given in [33], and it was used in [6, 7], where it was helpful in the feedback control implementation.

This problem contains nonlinear dynamics in the surface reactions, so we will use the nonlinear feedback tracking control method developed in [9], based on the solution to the state-dependent Riccati equation [34, 35, 36, 37, 38]. With this method the feedback control is given by

$$u(t, x^M) = -R^{-1}\mathcal{B}^T \left[\Pi(x^M)x^M + s(t, x_{nom}^M) \right], \quad (43)$$

where $\Pi(x^M)$ is the solution to the state-dependent Riccati equation

$$\Pi(x^M)\mathcal{A}(x^M) + \mathcal{A}^T(x^M)\Pi(x^M) - \Pi(x^M)\mathcal{B}R^{-1}\mathcal{B}^T\Pi(x^M) + (H^M)^T Q_1 H^M + (\bar{H}^M)^T Q_2 \bar{H}^M = 0. \quad (44)$$

The term $(\bar{H}^M)^T Q_2 \bar{H}^M$ comes from the cost functional weight on \bar{d}^M mentioned earlier. This SDRE is solved using a power series approximation described in [9, 36], beginning by rewriting the matrix \mathcal{A} as a sum of constant and state-dependent parts by $\mathcal{A}(x^M) = A_0 + \varepsilon g(x^M)\Delta A_C$ with a temporary variable ε . The SDRE solution $\Pi(x^M)$ is expanded as a power series in ε , resulting in

$$\Pi(x^M, \varepsilon) = \sum_{n=0}^{\infty} \varepsilon^n g^n(x^M)(L_n)_C. \quad (45)$$

Substituting these two expansions into equation (44) and matching terms with the same powers of ε results in the following set of equations for determining the constant-valued matrices $(L_n)_C$:

$$(L_0)_C A_0 + A_0^T (L_0)_C - (L_0)_C \mathcal{B}R^{-1}\mathcal{B}^T (L_0)_C + (H^M)^T Q_1 H^M + (\bar{H}^M)^T Q_2 \bar{H}^M = 0, \quad (46)$$

$$(L_1)_C \left[A_0 - \mathcal{B}R^{-1}\mathcal{B}^T (L_0)_C \right] + \left[A_0^T - (L_0)_C \mathcal{B}R^{-1}\mathcal{B}^T \right] (L_1)_C + (L_0)_C \Delta A_C + \Delta A_C^T (L_0)_C = 0, \quad (47)$$

$$(L_n)_C \left[A_0 - \mathcal{B}R^{-1}\mathcal{B}^T (L_0)_C \right] + \left[A_0^T - (L_0)_C \mathcal{B}R^{-1}\mathcal{B}^T \right] (L_n)_C + (L_{n-1})_C \Delta A_C + \Delta A_C^T (L_{n-1})_C - \sum_{k=1}^{n-1} \left[(L_k)_C \mathcal{B}R^{-1}\mathcal{B}^T (L_{n-k})_C \right] = 0. \quad (48)$$

The first N_P of these terms are then substituted into equation (45), with ε set to 1, to obtain an approximation to the SDRE solution. In our HPCVD problem the components of $\mathcal{A}(x^M)$ are

$$A_0 = \begin{bmatrix} A^{MG} & 0^{MG \times Ms} \\ H_q^{MG}/W_{Ga} & -k_1 & 0 & 0 \\ 0^{1 \times MG} & k_1 & -k_{GaP}S_p & 0 \\ 0^{1 \times MG} & 0 & k_{GaP}S_p & 0 \end{bmatrix}, \quad (49)$$

$$\Delta A_C = \begin{bmatrix} 0^{MG \times MG} & 0^{MG \times Ms} \\ 0 & 0 & 0 & 0 \\ 0^{Ms \times MG} & 0 & 0 & k_{GaP} \\ 0 & 0 & 0 & -k_{GaP} \end{bmatrix}, \quad (50)$$

with $g(x^M) = n_2$. In the simulations that follow we use $N_P = 5$ power series terms in the expansion for $\Pi(x^M)$ in equation (45).

The other part of the control formulation is the tracking variable $s(t, x_{nom}^M)$, which is found using a two-point boundary value problem technique developed in [9]. The tracking variable and a

nominal state variable form the coupled system

$$\begin{cases} \dot{s} &= -\mathcal{A}^T(x_{nom}^M)s + \Pi(x_{nom}^M)\mathcal{B}R^{-1}\mathcal{B}^Ts + (H^M)^TQ_1d_T - D_t\Pi(x_{nom}^M)x_{nom}^M \\ &\quad - \sum_{i=1}^M(x_{nom}^M)_i\left(\frac{\partial A_{1 \rightarrow M,i}}{\partial x_{nom}^M}(x_{nom}^M)\right)^T\left(\Pi(x_{nom}^M)x_{nom}^M + s\right) \\ \dot{x}_{nom}^M &= \mathcal{A}(x_{nom}^M)x_{nom}^M - \mathcal{B}R^{-1}\mathcal{B}^T(\Pi(x_{nom}^M)x_{nom}^M + s), \end{cases} \quad (51)$$

with initial condition $x_{nom}^M(0) = x_0^M$ and final condition $s(T_f, x_{nom}^M(T_f)) = 0$. The final time used in the two-point boundary value problem is given by the sum $T_f = T_p + \Delta$. Here T_p is the ending time of the studied target thickness trajectory, and Δ is an additional time value to allow this problem to approximate the infinite-time problem well and to be stabilized completely by the control. In a linear problem a reasonable value for Δ can be argued as five times the dominant time constant associated with the eigenvalues of the matrix $\mathcal{A} - \mathcal{B}R^{-1}\mathcal{B}^T\Pi$ (see p. 86 of [33]). For the nonlinear HPCVD problem, the eigenvalues of the linearized matrix $A_0 - \mathcal{B}R^{-1}\mathcal{B}^T(L_0)_C$ could be used. However, in addition to the nonlinearity, our desired tracking signal is a ramp function which is not zero at the target trajectory final time $T_p = 50$ (in dimensionless time units), which causes problems with forcing s to return to zero at only a short time after T_p . Thus we use a large value ($\Delta = 50$) to avoid this, although we are only concerned with the system behavior up to T_p .

The term $D_t\Pi(x)$ in the system (51) is a somewhat misleading notation; it is the total time derivative of $\Pi(x(t))$ given by

$$D_t\Pi(x) = \sum_{k=1}^m \frac{\partial \Pi}{\partial x_k}(x)\dot{x}_k, \quad (52)$$

and thus is only given meaning when evaluated along the state trajectory $x(t)$ so that \dot{x} can have some value. The terms $(\partial A_{1 \rightarrow m,i}/\partial x)$ represent the derivatives of the columns of A and are given by

$$\frac{\partial A_{1 \rightarrow m,i}}{\partial x} = \begin{bmatrix} \partial A_{1i}/\partial x_1 & \cdots & \partial A_{1i}/\partial x_m \\ \vdots & \ddots & \vdots \\ \partial A_{mi}/\partial x_1 & \cdots & \partial A_{mi}/\partial x_m \end{bmatrix}. \quad (53)$$

The numerical discretization of the two-point boundary value problem for s and x_{nom} on the interval $t \in [0, T_f]$ is carried out as follows. The x_{nom} variables have an initial condition, so a backward difference formula is applied to them, while the s variables have a final condition, so a forward difference formula is applied to them. This leads to the discrete system

$$\begin{cases} \frac{1}{\Delta t}(s_{k+1} - s_k) &= -\frac{1}{2}\left[A^T(x_k) - \Pi(x_k)\mathcal{B}R^{-1}\mathcal{B}^T\right](s_k + s_{k+1}) + \frac{1}{2}H^TQ(r_k + r_{k+1}) \\ &\quad - \sum_{i=1}^m(x_k)_i\left(\frac{\partial A_{1 \rightarrow m,i}}{\partial x}(x_k)\right)^T\left[\Pi(x_k)x_k + \frac{1}{2}(s_k + s_{k+1})\right] \\ &\quad - \frac{1}{\Delta t}[\Pi(x_{k+1}) - \Pi(x_k)]x_k \\ \frac{1}{\Delta t}(x_k - x_{k-1}) &= \frac{1}{4}\left[A(x_k) + A(x_{k-1}) - \mathcal{B}R^{-1}\mathcal{B}^T(\Pi(x_k) + \Pi(x_{k-1}))\right](x_k + x_{k-1}) \\ &\quad - \mathcal{B}R^{-1}\mathcal{B}^Ts_k, \end{cases} \quad (54)$$

where $s_k = s(k\Delta t)$ and $x_k = x_{nom}(k\Delta t)$, for the discretization points $k = 1, \dots, N-1$. Here N is the chosen number of discretization intervals, so that $\Delta t = T_f/N$. Also, x_0 is the given initial condition and $s_N = 0$ is the final condition. The x_{nom} variables are not averaged in the s equation, and the s variables are not averaged in the x_{nom} equation, because we have no values for x_N or s_0 . However, the discretization of $D_t\Pi(x_{nom})$ requires two time values to find the difference version

of the derivative, and thus includes an x_N value. For this we replace $[\Pi(x_N) - \Pi(x_{N-1})] / \Delta t$ by $[\Pi(x_{N-1}) - \Pi(x_{N-2})] / \Delta t$, making the assumption that this change will be small since near T_f the state will be stable and there will be very little change in x_{nom} there. The discretized system can be written in terms of the variables $y = [s_1, \dots, s_{N-1}, x_1, \dots, x_{N-1}]$ as $F(y) = 0$, resulting in a $2m(N-1)$ -dimensional system of nonlinear equations. This can be solved via the Newton method, with the next iterate found from the current one by

$$y_{n+1} = y_n - [\nabla F(y_n)]^{-1} F(y_n), \quad (55)$$

where ∇F denotes the Jacobian of F .

This tracking variable, found offline, is combined with the direct state feedback term involving the state-dependent Riccati equation solution in the feedback control formulation (43). The tracking variable is a very important part of the control for the HPCVD problem, since there is some delay between the introduction of controlled species at the inlet and the arrival of that species at the substrate surface (where the tracking signal d^M as well as the absorption measurement $z^M(x^M)$ are evaluated). The tracking variable anticipates this delay, whereas the direct state feedback term $-R^{-1}\mathcal{B}^T\Pi(x^M)x^M$ cannot do so.

The application of the above feedback control to the HPCVD system must also include a state estimator/compensator based on the measurement discussed in Section 2. The control described in equation (43) is given in terms of the full state variables x^M (the gas-phase POD coefficients and ROSK model species concentrations), but not all of these state variables are known. We have only a partial measurement, the optical absorption across the width of the reactor, with which to find an estimate x_e^M of the actual state. Rewritten in terms of the POD basis representation, the absorption measurement in equation (33) becomes

$$\begin{aligned} z^M(t) &= 1 - \exp\left(-\frac{4\pi W_0}{\lambda} \sum_{n=1}^3 \sum_{i=1}^{M_n} \frac{b_n}{W_n} \left[\int_{\mathcal{W}} \Phi_{ni}(\vec{x}) d\vec{x} \right] y_{ni}(t)\right) \\ &= 1 - \exp\left(-C_0^M x^M(t)\right) \\ &= c(x^M(t)). \end{aligned} \quad (56)$$

This measurement is nonlinear in the state, as are the dynamics in the system (40), so we will use the nonlinear techniques for state estimation developed in [9].

The estimated state will be represented by an ordinary differential equation similar to the state equation, with a gain matrix (found using a state-dependent Riccati equation) applied to the difference between the measurements of the actual and estimated states. With the feedback control (43) given in terms of the estimated state, the actual and estimated states are coupled and are given by

$$\begin{cases} \dot{x}^M &= \mathcal{A}(x^M)x^M - \mathcal{B}R^{-1}\mathcal{B}^T \left[\Pi(x_e^M)x_e^M + s(t, x_{nom}^M) \right] \\ \dot{x}_e^M &= \mathcal{A}(x_e^M)x_e^M - \mathcal{B}R^{-1}\mathcal{B}^T \left[\Pi(x_e^M)x_e^M + s(t, x_{nom}^M) \right] \\ &\quad + L^M(x_e^M) \left[z^M - c(x_e^M) \right]. \end{cases} \quad (57)$$

The state estimation gain is found by

$$L^M(x_e^M) = \Sigma(x_e^M)(C_0^M)^T V^{-1} \quad (58)$$

from the solution $\Sigma(x_e^M)$ to the dual state-dependent Riccati equation

$$\Sigma(x_e^M)\mathcal{A}^T(x_e^M) + \mathcal{A}(x_e^M)\Sigma(x_e^M) - \Sigma(x_e^M)(C_0^M)^T V^{-1} C_0^M \Sigma(x_e^M) + U = 0. \quad (59)$$

We set the weights in the estimation problem as $U = I_M$ and $V = r_3$, with r_3 a third design parameter we can choose in order to vary the emphasis in the control/estimator formulation. For the purposes of finding the estimator gain in equations (58)-(59), the nonlinear measurement function $z^M = c(x^M)$ is linearized about the origin, resulting in

$$\begin{aligned} z_0^M(t) &= c(0) + \frac{\partial c}{\partial x^M}(0)x^M(t) \\ &= C_0^M x^M(t). \end{aligned} \quad (60)$$

The nonlinearity of the measurement function does remain in the estimator system (57) itself, and the nonlinearity of the system dynamics remains in (57)-(59). The estimator gain SDRE (59) is solved using the power series approximation in an analogous fashion to equations (44)-(48) by splitting $\mathcal{A}(x^M)$ into constant and state-dependent parts as before and using $N_P = 5$ power series terms.

5. Film Thickness Tracking Results

The POD modes for each gas-phase species were found in the manner described in Section 2, from sets of 100 snapshots for each species. The percentage of data variability contained in the first k POD modes is plotted in Figure 5, for each of the three species. This shows that the original

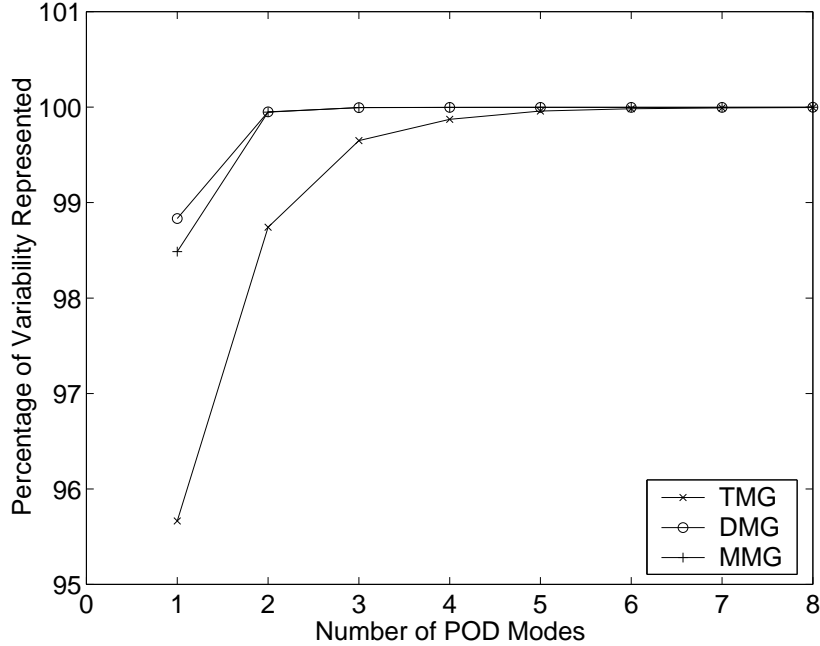


Figure 5: Data variability contained in the first few POD modes, for each species.

snapshot data is very well represented by 6 modes or fewer for the TMG mass fraction, and 3 modes

or fewer for the other two species, suggesting that no more than 12 POD modes are required for the reduced order model. This is in contrast with the 3×47131 finite element basis functions used in the reactor simulation from which the POD snapshots were derived.

However, while variability may be a good guide to the accuracy of the POD model as a representation of the full system, the effectiveness of the POD model as the basis of feedback control is another issue. For this purpose a study of the controllability and observability properties of the reduced order system may be helpful. As in [7, 39, 40] we may look at the controllability and observability ranks of the resulting linear POD reduced order representation (29) of the gas phase dynamics, with the linearized measurement function given in terms of the gas-phase states by

$$z_0^{M_G}(y^{M_G}) = C_0^{M_G} y^{M_G}. \quad (61)$$

Control theory results [41] yield that a linear system will become less controllable and/or observable as the difference between those ranks and the system rank increases. This has appeared to be so in some cases [39, 40] where a greater number of POD modes made the control more difficult, though in others [7] a larger POD system was found to produce a significantly more effective control than a smaller POD system.

For our system, the complete HPCVD model is a coupled system of the gas-phase and surface deposition models. While the complete model is nonlinear, the gas-phase POD system is linear, and so the above techniques of considering data variability and controllability/observability properties can be used in finding the number of POD modes. In the POD reduced order model the maximum controllability rank is 5, and the maximum observability rank is 6. In order to remain very close to the controllability and observability ranks in the POD system, we use $M_G = 6$ modes total (2 for TMG, 2 for DMG, and 2 for MMG) in constructing the feedback control and estimator. This low number of modes still results in the capture of a very large percentage of the data variability as shown in Figure 5. It should be noted that the heuristic arguments involving these ranks are for the reduced order system itself, but the eventual goal for a model of this type is to apply the reduced order control to the high-order system, or indeed the original infinite-dimensional system. It is questionable whether controllability in the full problem is even related to that of the reduced order system. Thus even though the controllability rank appears to be something which might be heuristically considered, strict adherence to any rule based on it need not produce optimal results.

In implementing this problem numerically, the system (9) is nondimensionalized before applying the Galerkin procedure to produce an approximated ODE system. We used reference values of $L_0 = 10^{-3}$ m, $D_0 = 10^{-5}$ m²/s, and $\rho_0 = 3.04$ kg/m³, with other scaling values extrapolated from them. The surface model is similarly nondimensionalized before the simulations are run. From the FIDAP-generated snapshots, the POD modes and the coefficient matrices A^{M_G} , B^{M_G} , $H_q^{M_G}$ and $C_0^{M_G}$ are calculated using C programs. The matrix calculations involved in setting up the control formulation are done in MATLAB programs, including the determination of Riccati and Lyapunov equation solutions in (46)-(48) using the MATLAB routines “*are*” and “*lyap*” respectively, and the solution of the TPBV problem for the tracking variable using codes written by the authors. The dynamical equations in the actual control problem are solved in MATLAB as well using the “*ode23s*” function. The initial estimated state $x_e^M(0) = \mathcal{O}(10^{-8})$ only approximates the actual state initial condition $x^M(0) = 0$, which represents a reactor empty of all but the carrier gas, and no species on the surface except the activated phosphorus, which is not one of the state variables. The slightly inaccurate initial estimated state is chosen to analyze the effectiveness of the state estimation method at recovering so as to accurately represent the actual state.

We considered the feedback control of the reduced order system using 6 POD modes, tracking a film thickness growth profile shaped as in Figure 4. The values of the design parameters $r_2 = 0$ and $r_3 = 4000$ were chosen because they gave the best results among a number of different values which were tried in the control formulation. A series of simulations was carried out with these values of r_2 and r_3 kept constant and the third parameter given the various values $r_1 = 10^{-9}$, 10^{-8} , and 10^{-7} . The resulting film thicknesses from the controls using these three parameter values are plotted in Figure 6, tracking the growth of one monolayer of GaP (with dimensionless time units equivalent to 0.1 s). This figure depicts excellent tracking of the desired thickness profile, with the trajectory

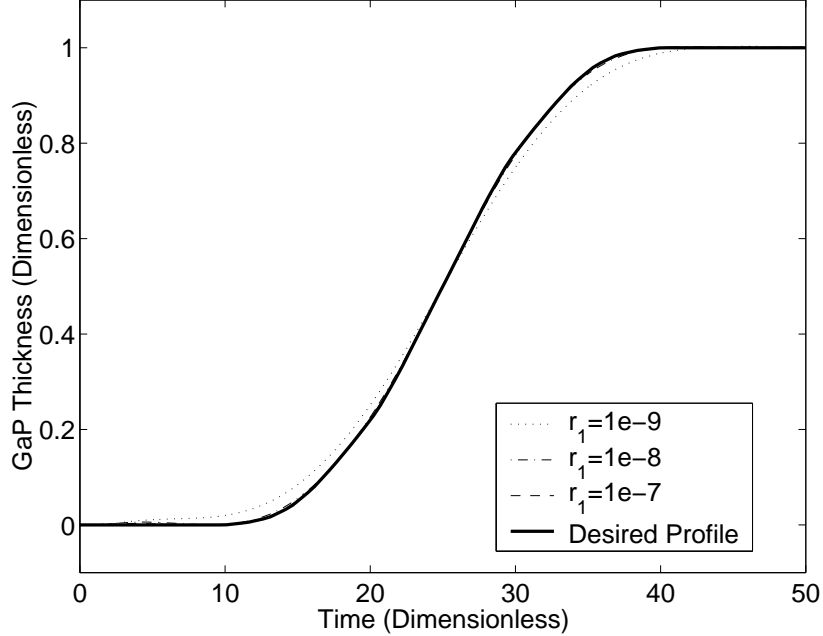


Figure 6: Controlled thickness profiles for various r_1 values.

moving closer to the desired curve as the value of r_1 for the thickness weighting term in the cost functional increases, although even the $r_1 = 10^{-9}$ output is very close to the target profile.

The control inputs u which were used to achieve these three controlled trajectories are plotted in Figure 7, with the (scaled down) desired trajectory plotted there as well to show relative dynamics between the control action and target trajectory. Each control here is a roughly pyramid-shaped, or hat-shaped, pulse. Each goes slightly negative at the beginning and end of the hat shape, and there is a very sharp spike in each, located immediately at the start of the time interval considered, before the hat shape itself starts. As r_1 increases, both the hat peak and the negative portions of the control become larger. The larger r_1 values (and so larger Q_1 and relatively smaller R in the cost functional) allow greater extremes in the control, which then force the thickness ramp to have sharper curves and more closely approach the rounded-ramp-shaped desired growth profile. The initial spike in the control input appears to be related to the slightly inaccurate initial estimated state which we used, although it may also be a very brief period in which the control/estimator needs to “learn” the behavior of the system.

Figure 8 shows the root-mean-square error between the nodal point values of the gas-phase mass

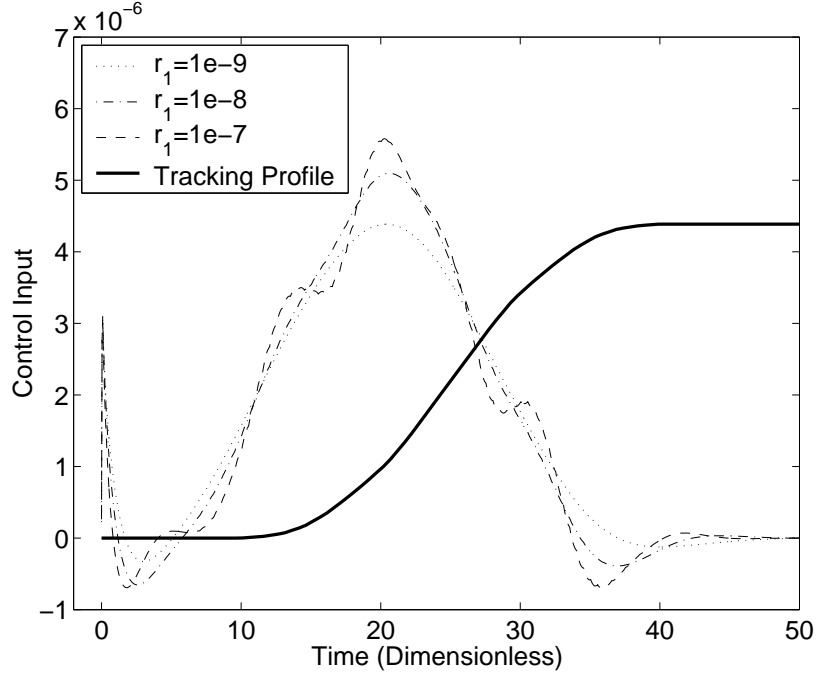


Figure 7: Control inputs for various r_1 values, with not-to-scale tracking profile shown.

fractions obtained from the actual and estimated states,

$$\text{RMS} = \sqrt{\frac{1}{3N} \left[\vec{Y}^{3N}(t) - \vec{Y}_e^{3N}(t) \right]^2}, \quad (62)$$

during the controlled simulations. Note that this error decreases extremely rapidly (the time interval shown in this figure is much smaller than that in the thickness and control figures), so that it is almost zero by the nondimensional time 5. This error curve is essentially the same for all three values of r_1 , and shows that the estimator is performing well at approximating the actual state from the nonlinear partial measurement. It also implies that the very sudden, very short control spike may be necessary to close the gap between the estimate and the actual state.

Other simulations were run with a different design parameter value of $r_3 = 1$. This much smaller value changes the weights in the state estimation cost functional so that more “control” is allowed to close the difference between the estimated and actual states more quickly. This results in the initial control spike being much larger, going from -11 to 7 , in comparison with the spike in Figure 7, which has a limited range of 0 to 3 . Increasing r_3 much beyond 4000 results in the system becoming badly scaled and “*ode23s*” being unable to solve it.

Without constraints on the possible values the control input can take, it is possible for the control to go negative, as can be seen in Figure 7. Physically, the control input represents the TMG mass fraction at the inlet. Therefore, having negative values of the control is undesirable, since it would be impossible to realize in the physical reactor. In a previous study [7], similar nonphysical negative control was truncated, or set to 0 whenever the control formulation gave a negative value. In those gas-phase-only simulations, the feedback control successfully tracked the surface flux profile, even

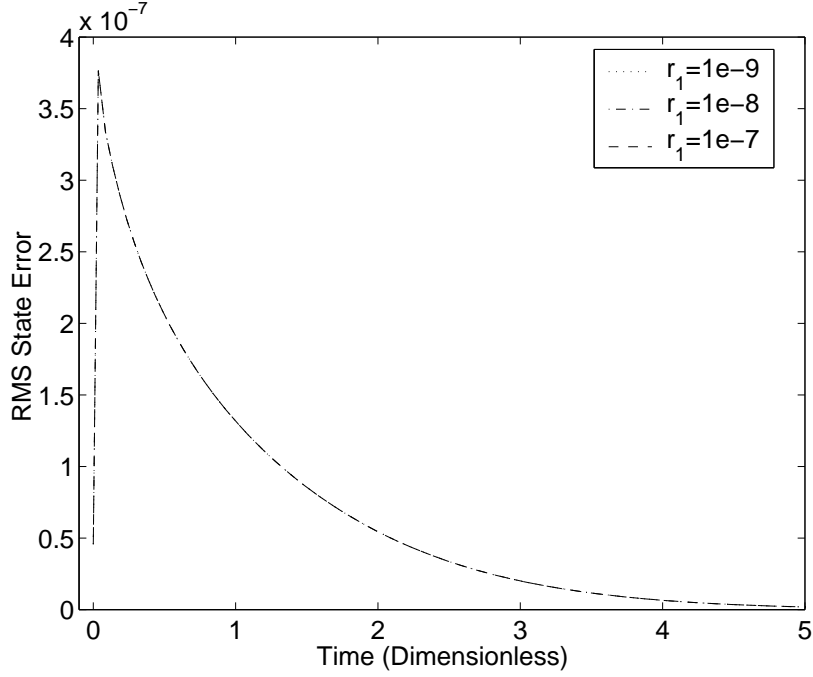


Figure 8: State estimation error amounts for various r_1 values.

with this restriction. In the control problems in this paper we attempted to use the same truncation technique to eliminate negative control inputs, but found that with truncation this problem would not converge to the desired result. There are a few aspects of the HPCVD problem which can help to explain this ineffectiveness.

An important property of the film thickness tracking problem which contributes to the failure of truncation is that the tracking signal (film thickness) is a cumulative measure of the amount of GaP which has been grown during the cycle up to the current time. In [7], where truncated control was successfully implemented, the tracking signal was the flux of gallium to the surface, which is an instantaneous property of the system. In that case, if a negative control input was truncated near the end of the desired block-shaped flux profile, it would simply lead to a slight trailing over of the gallium flux past the desired stopping time, but no long-term unwanted effects. In fact the target profile in [7] includes a slight trailing end in anticipation of this effect. In contrast, if no negative control is allowed in our current problem, any excess active gallium on the surface which the negative control would have removed instead goes into growing more of the GaP film, thereby overshooting the upper plateau of the target ramp function. The film thickness would then differ from the desired value for the rest of the time period instead of just a short interval. This adds substantially to the cost functional. This is the primary reason for the failure of the truncated feedback control. If the HPCVD model was extended to consider long-term film growth this problem might be removed, since any overshoot could be compensated for with less growth in the next source vapor pulse cycle. The next phosphorus source pulse could be used to control the leftover active surface gallium through reactions in the next growth period.

It is also important to point out that the target profile shape profoundly affects the system

behavior. Our target thickness profile is a upward ramp with rounded corners. If the corners were not rounded at all, it would be virtually impossible to track closely, since sharp corners are extremely difficult to control, especially in this case. In the surface reaction layer, even with a fairly quick reaction rate for the formation of GaP, there is still some time needed for the reaction to take place (likewise for the activation of the surface gallium). Since the rate of formation is proportional to the amount of active gallium, as it is used up, the reaction will go slower, prolonging and slowing the end of the film growth process. In order to cause a sudden end to the GaP growth, the feedback control sends negative gallium concentrations out to “eliminate” the gallium remaining on the surface at the end of the desired growth period. (We note that in discussions with our material scientist colleagues, the nature of this control has already suggested an enhanced feature of the next generation of reactors. It is proposed to inject a third chemical species to combine with the gallium, producing an inactive, nondepositing product, thereby producing timely “sinks” or negative control on the gallium.) This difficulty with negative control is naturally lessened when the corners of the desired thickness ramp function are smoothed out. With the profile shown in Figures 6 and 7 there is only minimal negative control, especially with the parameter $r_1 = 10^{-9}$. There is a less dramatic effect in the gas phase, where any sharp jump in the gallium input at the inlet (such as a block-shaped pulse) will be smoothed out by gas flow/diffusion in the reactor before it arrives at the surface, making sharp start/stop points difficult. This was compensated for by using a somewhat rounded target profile in the gas-phase simulations in [7], and should be compensated for in the current simulations by the rounded target thickness ramp function as well.

A different set of simulations using a more ambitious, less rounded target profile was carried out as well. In this profile the ramp function is less smoothed out, with a growth period of 20 time units compared to 30 in the previous simulations. All the design parameters are the same as above except the time constants $T_p = \Delta = 30$, and the parameter r_3 . For these simulations the value $r_3 = 1000$ gave the best results. The thickness profile results for these simulations using the same three values of r_1 as before are plotted in Figure 9. The corresponding control inputs are plotted in Figure 10 along with the not-to-scale target thickness trajectory. We can track this sharper thickness profile almost as well as the earlier profile, but with greater difficulty in keeping the control nonnegative. Notice that in Figure 10 the $r_1 = 10^{-7}$ control input goes down to almost -2 as the growth is levelling off, while in Figure 7 it only reaches about $-2/3$. Even the $r_1 = 10^{-9}$ control input goes down to nearly -1 in the initial spike in Figure 10, while only going to about $-1/3$ near the start in Figure 7. It might be possible to remove the negative control inputs entirely with enough rounding and slowing of the desired film growth. However, this would limit the rate at which the film could be grown. To determine the proper balance between the speed of growth and the ability to control the process, so that we can find the target profile that is most desirable yet can be realistically achieved, we will need more information about the HPCVD reactor and model (accurate values of the parameters k_1 and k_{GaP} , etc). The estimator error in the sharper-profile simulations is reduced to very near zero by about time 1.5, even earlier than the time 5 needed by the error in Figure 8. This is a result of the parameter $r_3 = 1000$ in the estimator weights being smaller than the Figure 8 value of $r_3 = 4000$.

It can be noted that in all of the above simulations we have used the value $r_2 = 0$ for the added state weight term on \bar{d}^M in the cost functional. Simulations were run with several positive values of r_2 , but using this did not seem to be beneficial overall. There did not appear to be any significant undesirable behavior in the state variables which needed to be limited by this term, and diverting the feedback control influence to include the other states reduced the ability of the control to achieve

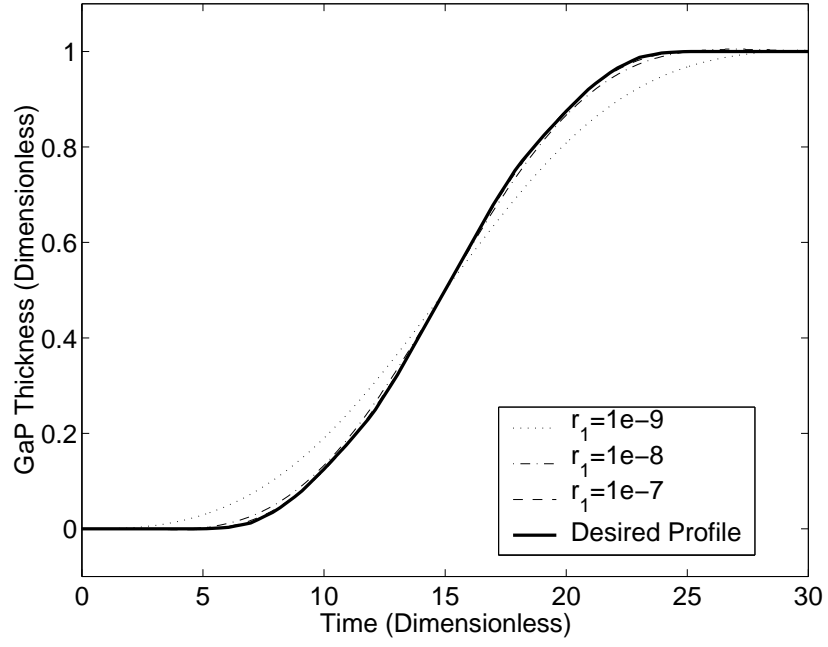


Figure 9: Controlled thickness profiles with sharper target profile.

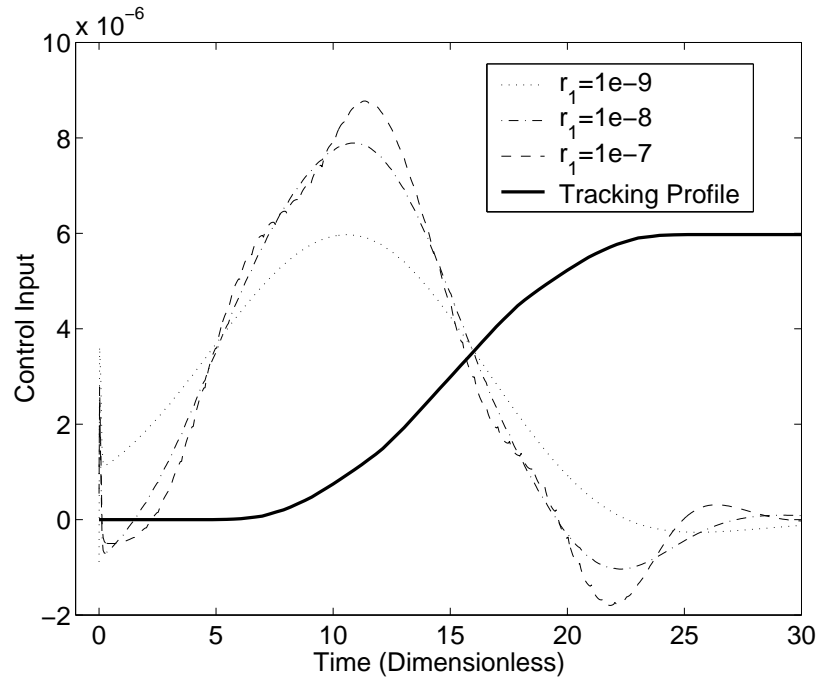


Figure 10: Control inputs with sharper target profile (not-to-scale tracking profile shown).

the desired tracking results. Simulations using $r_2 = 10^{-9}$ and other positive values were significantly less effective at tracking the desired thickness profile than those using $r_2 = 0$, and the initial control spike was much more negative as well.

6. Discussion

While we have successfully carried out proof-of-concept model development and development of a nonlinear control methodology for a very complex reactor system, there are several directions in which this research work should be continued/expanded on in the future. We have developed a reduced order model-based feedback control and implemented it on the reduced order model of the HPCVD problem. However, the success of this is not necessarily indicative of a success upon application of the reduced order control to the full system. This should be tested by implementing the reduced order control on the high-order finite element representation, as was done for the gas-phase HPCVD model in [7]. Furthermore, the ultimate goal is to apply the reduced order control to the physical reactor, which has recently been constructed in the Materials Science Department at North Carolina State University.

In order to make the model sufficiently accurate to effectively control the physical reactor, some refinements will be necessary. Several parameters we have used here are very rough estimates, used to prove the effectiveness of the control on a problem with behavior very similar to that of the physical problem. For the actual reactor, however, we will need to find more accurate values for the ROSK model rate constants and the optical responses for the gas-phase species, through calibration of and experiments in the reactor. In addition, there will be a second observation technique (a modification of the PRS sensing of [4, 8]) added in the reactor, using the measurement of the reflectance of a light beam from behind the substrate. The equations modeling this must be added to the state estimation portion of the control in order to use this measurement in controlling growth and composition of the film in the physical reactor.

A different avenue of investigation is the number of POD modes used in the reduced order model. We have used 6 POD modes, but using a larger number of modes (perhaps 10 or 12) may result in a more accurate representation of the actual three-dimensional species transport dynamics, and therefore possibly a more effective model-based control. Application of the 6 mode model, as well these larger models, to control of the high order finite element model would be very useful in determining which of them should give us the best control authority over the physical reactor. The properties of the nonlinear control and estimation methodologies used here could also be further investigated computationally, to learn more about conditions that result in robustness and local and global stability, for example.

Another important point is that in this proof-of-concept presentation we have considered only a single cycle of the deposition process, concentrating on the small-time-scale behavior of the system. While this is important, controlling the long-term growth behavior over a period of hundreds of cycles is another important goal. This will mean extending the model to include the gas-phase phosphine transport and decomposition in addition to that of TMG, and changing the ROSK model to include the surface phosphorus dynamics. Adding phosphine to the gas-phase model may well require the removal of the dilute assumption described in Section 2, since phosphine has a very small sticking coefficient and so large amounts of it must be introduced in order to get the desired amount deposited on the substrate surface. Some modifications to the control problem formulation may prove necessary in this event as well. A benefit of extending the control problem over many cycles

would be that the problem with negative controls might be removed, since overshooting a target film thickness might be compensated for in the next source vapor pulse cycle. The addition of the second measurement technique should also be helpful in observation and control of the film growth behavior over many cycles, since the optical absorption measurement focuses only on the gas phase while the reflectance measurement would provide a more direct observation of the cumulative film thickness growth. A further possibility is that the problem could be extended to look at the growth of films such as $\text{Ga}_{1-x}\text{In}_x\text{P}$, where both the film thickness and film composition can be controlled, by adding more species and reactions to both the gas-phase and surface models.

Finally, we note that the present investigation opens a host of mathematical questions that should provide a rich source of theoretical challenges. For example, there is no general (or even particular) theory for how one chooses snapshots in the POD methodology. This is true for general simulation approximations, but even more relevant in specific constructions for control and inverse problems. That is, does one use snapshots of the uncontrolled system in constructing the basis elements to be used in the controlled (feedback) system? In most cases this will not prove adequate. Then how does one systematically choose “inputs” for the system to be used in taking snapshots? These are fundamental questions of both practical and theoretical interest. Successful treatment will almost certainly require rather deep and detailed mathematical analysis combining dynamical systems and control theoretical ideas.

Very little in the way of a theoretical convergence (rigorous error estimate) analysis is available for POD-based approximations (we are aware of only the recent efforts of Kunisch and Volkwein in [42]). Again this is an important question that will require nontrivial mathematical analysis.

Another area of significance is the development of a comprehensive theory for approximation and convergence for the state-dependent Riccati equation approach to nonlinear feedback and compensation that we have adapted for our computations above. At present, there is no rigorous theory that treats these approximation ideas for complex systems such as the nonlinear coupled ordinary/partial differential equation system of our reactor model. The development of a rigorous foundation for such methods is a formidable challenge to the control theory community. Even partial theories will provide insight into further development and use of nonlinear control methodologies.

7. Conclusions

In this paper we have developed a mathematical model of the thin film growth process in a high-pressure chemical vapor deposition reactor. This model combines the gas-phase flow of the source species to the substrate with the growth of the film on that substrate surface. The film growth is represented by a reduced order surface kinetics model, and the species transport is represented by a reduced order model obtained from a quasi-steady gas flow formulation and the proper orthogonal decomposition method. This gas-phase model includes a three-dimensional domain, species reactions, and a realistic technique for partial measurement of the process. The full HPCVD model is formed by linking these two parts of the growth process through the gas-phase species flux to the substrate surface.

The low order models representing the two parts of the film deposition process enable us to implement a real-time model-based nonlinear feedback control of the film growth. A feedback control tracking GaP film thickness was constructed for this problem, using a nonlinear partial measurement for estimation of the state. Previously developed methods of nonlinear tracking control and state estimation based on two state-dependent Riccati equations were used for this purpose. In

simulations the inlet TMG mass fraction was used to effectively control the model and track a desired film thickness growth profile, with the estimator approximating the actual state very well, although there was the drawback of some slightly negative control input values near the film growth starting and stopping times. (This, in fact, stimulated new ideas for future generation reactor design as mentioned in Section 5.) The importance of choosing the right target film growth profile was discovered; trying to force the growth to be too fast, or to start and stop too sharply, resulted in large negative control input portions. On the other hand, choosing too slow of a target growth profile will result in the loss of one of the advantages of the HPCVD process: its high throughput in comparison with low-pressure methods. Several future extensions of this work were discussed in Section 6, including modifications of this particular model and theoretical issues regarding the mathematical techniques used here.

8. Acknowledgements

This research was supported in part by a DOD/AFOSR MURI Grant AFOSR F49620-95-1-0447, in part by the AFOSR under Grant F49620-96-1-0292 (AASERT), and in part by the AFOSR under Grant F49620-98-1-0180. Part of the research was completed while one author (H. T. Banks) was a visitor at the Institute of Mathematical Sciences, The Chinese University of Hong Kong.

References

- [1] Gevelber, M., Toledo-Quinones, M., and Bufano, M., "Towards Closed-Loop Control of CVD Coating Microstructures," *Materials Science and Engineering A*, Vol. 209, pp. 377-383 (1996).
- [2] Zhou, J. J., Li, Y., Pacheco, D., Lee, H. P., and Liu, X., "Virtual Control Simulator for Closed-Loop Epitaxial Growth," *Journal of Crystal Growth*, Vol. 175, pp. 52-60 (1997).
- [3] Warnick, S. C., and Dahleh, M. A., "Feedback Control of MOCVD Growth of Submicron Compound Semiconductor Films," *IEEE Transactions on Control Systems Technology*, Vol. 6, pp. 62-71 (1998).
- [4] Dietz, N., Woods, V., Ito, K., and Lauko, I., "Real-Time Optical Control of $\text{Ga}_{1-x}\text{In}_x\text{P}$ Film Growth by p -Polarized Reflectance," *Journal of Vacuum Science and Technology A*, Vol. 17, pp. 1300-1306 (1999).
- [5] Ly, H. V., and Tran, H. T., "Proper Orthogonal Decomposition for Flow Calculations and Optimal Control in a Horizontal CVD Reactor," CRSC Technical Report CRSC-TR98-13, N.C. State University (1998), and *Quarterly of Applied Mathematics*, to appear.
- [6] Kepler, G. M., Tran, H. T., and Banks, H. T., "Reduced Order Model Compensator Control of Species Transport in a CVD Reactor," CRSC Technical Report CRSC-TR99-15, N.C. State University (1999), and *Optimal Control Applications and Methods*, to appear.
- [7] Kepler, G. M., Tran, H. T., and Banks, H. T., "Compensator Control for Chemical Vapor Deposition Film Growth Using Reduced Order Design Models," CRSC Technical Report CRSC-TR99-41, N.C. State University (1999), and *IEEE Transactions on Semiconductors*, to appear.
- [8] Beeler, S., Tran, H. T., and Dietz, N., "Representation of GaP Formation by a Reduced Order Surface Kinetics Model Using p -Polarized Reflectance Measurements," *Journal of Applied Physics*, Vol 86, pp. 674-682 (1999).
- [9] Beeler, S., Tran, H. T., and Banks, H. T., "State Estimation and Tracking Control of Nonlinear Dynamical Systems," CRSC Technical Report CRSC-TR00-19, N.C. State University (2000).
- [10] Bird, R. B., Stewart, W. E., and Lightfoot, E. N., *Transport Phenomena*, New York, New York: John Wiley and Sons, 1960.
- [11] Slattery, J. C., *Momentum, Energy, and Mass Transfer in Continua*, New York, New York: McGraw-Hill, 1972.
- [12] *FIDAP (Fluid Dynamics Analysis Package) 7.0 Theory Manual*, Evanston, Illinois: Fluid Dynamics International, 1993.
- [13] Shadid, J. N., Moffat, H. K., Hutchinson, S. A., Hennigan, G. L., Devine, K. D., and Salinger, A. G., *MPSalsa: A Finite Element Computer Program for Reacting Flow Problems. Part 1 – Theoretical Development* (Sandia National Laboratories Report), Springfield, Virginia: National Technical Information Service, 1996.

- [14] Svehla, R. A., NASA Technical Report R-132, 1962.
- [15] Bayazitoglu, Y., and Ozisik, M. N., *Elements of Mass Heat Transfer*, New York, New York: McGraw Hill, 1988.
- [16] Lide, D. R., and Kehiaian, H. V., *CRC Handbook of Thermophysical and Thermochemical Data*, Boca Raton, Florida: CRC Press, 1994.
- [17] Shah, R. K., and London, A. L., *Laminar Flow Forced Convection in Ducts*, New York, New York: Academic Press, 1978.
- [18] Buchan, N. I., and Jasinski, J. M., "Calculation of Unimolecular Rate Constants for Common Metalorganic Vapor-Phase Epitaxy Precursors via RRKM Theory," *Journal of Crystal Growth*, Vol 106, pp. 227-238 (1990).
- [19] Larsen, C. A., Buchan, N. I., Li, S. H., and Stringfellow, G. B., "Decomposition Mechanisms of Trimethylgallium," *Journal of Crystal Growth*, Vol 102, pp. 103-116 (1990).
- [20] Hirschfelder, J. P., Curtis, C. F., and Bird, R. B., *Molecular Theory of Gases and Liquids*, New York, New York: Wiley and Sons, 1954.
- [21] Babuška, I., "The Finite Element Method With Penalty," *Mathematics of Computation*, Vol 27, pp. 221-228 (1973).
- [22] Barrett, J. W., and Elliot, C. M., "Finite Element Approximation of the Dirichlet Problem Using the Boundary Penalty Method," *Numerische Mathematik*, Vol 49, pp. 343-366 (1986).
- [23] Karhunen, K., "Zur Spektral Theorie Stochastischer Prozesse," *Annales Academiae Scientiarum Fennicae, Series A1, Mathematica Physica*, Vol 37 (1946).
- [24] Loève, M., "Functions Aleatoire de Second Ordre," *Compte Rend. Acad. Sci. (Paris)*, Vol 220 (1945).
- [25] Jackson, J. E., *A User's Guide to Principal Components*, New York, New York: John Wiley and Sons, 1991.
- [26] Fukunaga, K., *Introduction to Statistical Pattern Recognition*, New York, New York: Academic Press, 1972.
- [27] Lumley, J. L., *Stochastic Tools in Turbulence*, New York, New York: Academic Press, 1970.
- [28] Berkooz, G., *Turbulence, Coherent Structures, and Low-Dimensional Models*, Ph.D. dissertation, Cornell University, 1991.
- [29] Berkooz, G., "Observations on the Proper Orthogonal Decomposition," in Gatski, T. B., Sarkar, S., and Speziale, C. G., eds., *Studies in Turbulence*, New York, New York: Springer-Verlag, 1992, pp. 229-247.
- [30] Holmes, P. J., Lumley, J. L., Berkooz, G., Mattingly, J. C., and Wittenberg, R. W., "Low-Dimensional Models of Coherent Structures in Turbulence," *Physics Reports – Review Section of Physics Letters*, Vol 287, pp. 338-384 (1997).

- [31] Iollo, A., Lanteri, S., and Désidéri, J. A., “Stability Properties of POD-Galerkin Approximations for the Compressible Navier-Stokes Equations,” *Theoretical and Computational Fluid Dynamics*, Vol 13, pp. 377-396 (2000).
- [32] Theodoropoulou, A., Adomaitis, R. A., and Zafiriou, E., “Model Reduction for Optimization of Rapid Thermal Chemical Vapor Deposition Systems,” *IEEE Transactions on Semiconductor Manufacturing*, Vol 11, pp. 85-98 (1998).
- [33] Anderson, B. D. O., and Moore, J. B., *Optimal Control: Linear Quadratic Methods*, Englewood Cliffs, New Jersey: Prentice-Hall, 1990.
- [34] Pearson, J. D., “Approximation Methods in Optimal Control,” *Journal of Electronics and Control*, Vol 13, pp. 453-465 (1962).
- [35] Burghart, J. A., “A Technique for Suboptimal Control of Nonlinear Systems,” *IEEE Transactions on Automatic Control*, Vol 14, pp. 530-533 (1969).
- [36] Wernli, A. and Cook, G., “Suboptimal Control for the Nonlinear Quadratic Regulator Problem,” *Automatica*, Vol 11, pp. 75-84 (1975).
- [37] Krikelis, N. J. and Kiriakidis, K. I., “Optimal Feedback Control of Non-linear Systems,” *International Journal of Systems Science*, Vol 23, pp. 2141-2153 (1992).
- [38] Cloutier, J. R., D’Souza, C. N., and Mracek, C. P., “Nonlinear Regulation and Nonlinear H_∞ Control Via the State-Dependent Riccati Equation Technique: Part 1. Theory,” *Proceedings of the First International Conference on Nonlinear Problems in Aviation and Aerospace*, Daytona Beach, FL, May 1996.
- [39] Banks, H. T., del Rosario, R. C. H., and Smith, R. C., “Reduced Order Model Feedback Control Design: Computational Studies for Thin Cylindrical Shells,” *CRSC Technical Report CRSC-TR98-25*, N.C. State University (1998).
- [40] Banks, H. T., del Rosario, R. C. H., and Smith, R. C., “Reduced Order Model Feedback Control Design: Numerical Implementation in a Thin Shell Model,” *CRSC Technical Report CRSC-TR98-27*, N.C. State University (1998), and *IEEE Transactions on Automatic Control*, to appear.
- [41] Brogan, W. L., *Modern Control Theory*, 3rd ed., Englewood Cliffs, New Jersey: Prentice Hall, 1991.
- [42] Kunisch, K. and Volkwein, S., “Galerkin Proper Orthogonal Decomposition Methods for Parabolic Problems,” *Optimierung und Kontrolle*, Bericht Nr. 171, Universitat Graz, Austria (1999).